Journal of Evidence-Based Psychotherapies,
Vol. 19, No. 2, September 2019, 1-14.

FACTORS TO CONSIDER IN THE INTERPRETATION OF TRIAL RESULTS IN PSYCHOTHERAPY FOR DEPRESSION IN ADULTS, WITH REFERENCE TO SUICIDAL BEHAVIOUR

Valerie FLETCHER [1], *Vaughan EVANS** [2]

[1] MA, School of Health & Community Studies, Leeds Beckett University. Leeds, UK. LS1 3HE. v.fletcher@leedsbeckett.ac.uk; https://orcid.org/0000-0002-5686-3468

[2] MSc, CTSU, Richard Doll Building, Old Road Campus, University of Oxford. Oxford, UK. OX3 7LF, vaughan.evans@ndph.ox.ac.uk; https://orcid.org/0000-0002-2366-2301

Abstract

The results of randomised controlled trials and meta-analyses of trials in medicine have led to major improvements in treatments, and thereby significant reductions in mortality and morbidity worldwide. The implications of specific aspects of trial methodology for evaluating the effectiveness of psychotherapy for depression are explored, with reference to trial size, quality, bias, eligibility criteria and the use of non-inferiority design. In contrast to the treatment of systemic disease, pill placebo control in trials of psychotherapy for depression is not 'no treatment', and the use of such can reduce measured treatment effect.

Trials have entry criteria, and thus use a selective sample, findings may therefore not be generalizable. For example, in psychotherapy, specific trial exclusion criteria (suicidal behaviour) appear to raise a concern that a therapy may be widely adopted and include the treatment of a category of individual excluded from a practice-informing trial.

A diagnosis of depression is associated with a higher mortality, notably suicide, and the inter-relationship between depression and suicidal behaviour is explored. Aside from the direct utilisation of specific psychotherapies for suicidal behaviour; the effective treatment of depression, notably major depression, by psychotherapy, may lead to a reduction in the prevalence of suicidal behaviour.

Keywords: psychotherapy, depression, suicide, randomised controlled trial, meta-analysis, non-inferiority trial, subgroup analysis

* Correspondence concerning this article should be addressed to Valerie Fletcher
E-mail: V.Fletcher@leedsbeckett.ac.uk

Introduction

The results from randomised controlled trials (RCTs) have led to changes in medical practice since the landmark trial of the antibiotic streptomycin as a treatment for tuberculosis (Medical Research Council, 1948). The knowledge gained from this trial, and subsequent studies, led to reported 100 per cent cure rates for pulmonary tuberculosis, a disease which had killed half the patients who developed it, and helped establish the importance of randomised trials to medical research (Crofton, 2006). The MRC trial is credited with setting the standard for the conduct and reporting of modern randomised trials and 1948 has become established as the earliest year used by the Cochrane Collaboration for searching for RCTs (McDonald, Westby, Clarke & Lefebvre, 2002). Although since the 1948 MRC trial the number of reported randomized trials with 100 or more participants has shown a continuing increase (McDonald et al, 2002), trials may still lack statistical power due to inadequate sample size (insufficient numbers randomised). Statistically significant results with a relatively small sample size are possible where the treatment effect is *large,* however the reliable detection of *small or moderate* differences in health outcomes that arise from treatment, requires large-scale randomised evidence (Baigent, 1997).

A meta-analysis (MA) of randomised controlled trials is a statistical technique for combining the results of more than one trial (where the treatments being compared are the same in each trial) so that the combined result can provide greater statistical power (more randomised data) and thus may reliably enable the detection of *small or moderate differences* in outcome that may not otherwise have been found (due to the play of chance) in the individual trial results (Baigent, 1997).

Evidence from randomised controlled trials, and meta-analyses of trials, on the effectiveness of psychotherapy – principally counselling, cognitive behavioural therapy (CBT) and behavioural activation (BA) – is reviewed within the context of the recognised strengths and limitations of this study methodology. Specific aspects of trial and meta-analysis methodology, as applied to the evaluation of psychotherapy, are discussed in terms of the how results can be influenced, and thus how they should be interpreted.

Evidence for the interrelationship between depression and suicide is explored, and thus the role of psychotherapy in the treatment of the former, and the prevention of the latter.

Depression in Adults

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5, 2018) identifies symptoms to use as diagnosing tools based on combination and duration

including diminished interest in activities, feelings of worthlessness or excessive guilt, recurrent thoughts of death including suicidal thoughts.

The National Institute for Health Care Excellence in the UK (NICE) develops treatment guidelines used by the National Health Service (NHS) in England and Wales and provide a single source of reference for professionals, the public and patients (NICE, 2018). The guidelines 'Depression in adults: recognition and management' were developed (as with other guidelines) from the review of trial results and systematic reviews (NICE, 2009, updated 2018). The guidelines outline depression as a "broad and heterogeneous diagnoses", "central to it being depressed mood and/or loss of pleasure in most activities" and consider associated symptoms such as disturbed sleep, appetite changes, loss of energy, agitation or slow movements, poor concentration, indecisiveness, worthlessness feelings or suicidal thoughts.

In addition to the NICE guidelines, a recent development in the UK has been the Person-Centred Experiential Counselling for Depression (PCE-CfD) defining a distinct approach, using a model-specific, rather than the generic term of 'counselling', which may address identified concerns with the 2009 NICE guidelines (Barkham, Moller & Pybis, 2017).

Depression and Health

Individuals with a diagnosis of depression have a higher mortality than the general population, with a reported range of reduction of life expectancy, for a single depressive episode or recurrent depressive disorder, of 7 – 11 years (UK data in Chesney, Goodwin & Fazel, 2014). The association of depression with factors established as detrimental to health, such as poor diet, lack of exercise, smoking and excessive alcohol use, may account for the higher rate of mortality (Gilman et al., 2017), and thus, effective interventions to reduce depression will reduce associated health factors.

Mental Health and Risk of Suicide

In the results of a meta-analytic review of studies, a considerably greater suicide mortality rate for major mental disorders (including schizophrenia, depression and bipolar disorder) was reported than that for the general population (Chesney et al., 2014). A retrospective case note study (Pearson et al., 2009) on individuals in the UK who had committed suicide, found that around 90% had consulted their General Practitioner (GP) on at least one occasion in the year before death; in over half the cases reviewed the primary reason for the final consultation was psychological symptoms. The report indicated that, in terms of risk assessment, agreement between primary (GP), and secondary care (specialist mental health teams), was poor (Pearson et al., 2009).

Depression and Risk of Suicide

The association of the specific mental illness of depression with suicide, or suicide risk, is affirmed in the results a range of large-data studies from around the world:

- A major depressive episode was described as the "most important risk factor for lifetime suicide attempt", with a population attributable risk proportion (PAR) of approximately 28%, in the results of a large pan-European study of risk factors for suicidality (Bernal et al., 2007, p. 31).
- In the USA, depression is one of the most common conditions treated in primary care; in patients referred to secondary care following a suicide attempt, around 60% were found to have depressive disorders (Dhossche, Meloukheia & Chakravorty, 2000).
- A markedly higher suicide rate in those with a diagnosis of depression, compared to the general population, has been reported (in the results of a meta-review) as around twenty times higher (standardised mortality ratio 19.7; 95% confidence interval 12.2 - 32.0) (Chesney et al., 2014).
- Around 15-20% of depressive patients have been reported as committing suicide (Goodwin & Jamison, 1999, in Miret et al., 2013).

Suicide Prevalence

Total annual deaths by suicide amount to over 800,000 people worldwide, with an estimate that for each completed suicide, more than 20 attempt suicide (WHO, 2014). The worldwide total will reflect country-by-country variations in prevalence, and the accuracy of each country's death registrations, and is probably under-reported and/or incorrectly classified in some countries (WHO, 2014). In Great Britain there were 5,668 suicides in 2016, around three-quarters male and the most common method was hanging (Office of National Statistics, 2017). In England and Wales in 2016, deaths by suicide and injury or poisoning of undetermined intent, at ages 20 to 34 years, remained the leading cause of death for both males and females in this age group, and the leading cause of death for males aged 35 to 49 years (ONS, 2017).

Suicide Risk Reduction

Suicidal behaviour encompasses completed suicide, suicide attempts, suicidal intent, plans and suicide ideation (Schneider, 2012). No single approach to suicide prevention has been reliably established as superior to any other due to a low number of RCTs, and the need has been identified to evaluate "with robust research designs" *combinations* of prevention strategies (Zalsman et al., 2016, p. 646). Strategies with evidence of efficacy include "restricting access to lethal means" (analgesics; known locations for suicide by jumping), school-based awareness programmes and pharmacotherapy (Clozapine, Lithium) (Zalsman et al., 2016, p. 646).

The efficacy of the most-widely used psychotherapeutic interventions (CBT and Dialectical Behavioral Therapy, DBT) to reduce suicide risk (suicidal ideation or suicide attempts) is supported by the results of observational studies, however with considerable variation in the results (Méndez-Bustos, et al., 2019). Evidence from randomised trials supports the use of CBT specifically "focussing on suicidal cognitions and behaviors" (Mewton & Andrews, 2016, p. 21).

Irrespective of the mode of intervention, intensive outpatient support therapy is a "mainstay" of suicide prevention guidelines (Méndez-Bustos et al., 2019, p. 2).

Trials in Psychotherapy

The use of reliable, evidence-based, psychotherapy interventions for depression can reduce the prevalence of this condition, and also lead to a reduction of the prevalence of suicidal behaviour; Bernal et al (2007, p. 31) noted that "major depressive episode appeared to be the most important risk factor for lifetime suicide attempt" (in their study sample) and therefore that the prevention of such (major depressive episodes) could reduce the "lifetime prevalence of suicide attempts" by nearly a third.

Modern evidence-based practice *in all fields of medicine* is informed by the results of RCTs (and meta-analyses of trials); there are, however, *specific aspects* of the methodology that necessitate consideration when evaluating results in trials of psychotherapy.

Factors which may Limit the Applicability of Trial Results for Informing Practice

Sample Size

A study with low statistical power not only has a diminished probability of determining a real treatment effect, but also the low power "…negatively affects the likelihood that a nominally statistically significant finding actually reflects a true effect." (Button et al., 2013, p. 365).

Cooper (2011, p. 4) noted, in a Briefing Paper for the British Association for Counselling and Psychotherapy (BACP), that "only a small number of RCTs of counselling for depression" existed, and that counselling-related RCTs tended to have "comparatively low numbers" randomised, and thus a low statistical power to reliably determine significant results. Some more recent trials in psychotherapy for depression have been of a larger scale, randomising several hundred participants. For example, the COBRA trial, described as the "largest in BA", and "one of the largest trials of psychological treatments for depression", randomised 440 participants (Richards et al, 2016, p. 879). The PRaCTICED trial compares Counselling for Depression versus CBT, and is described as a "large-scale" study,

with a planned recruitment of 550 (Saxon et al., 2017); the CADET study randomised 581, and compared the effectiveness in a primary care setting of collaborative care versus usual care, in a cluster RCT (Richards et al., 2013).

In medicine generally, large-scale trials, however, may randomise thousands of patients, with some exceeding 10,000 (ISIS-2, 1998, ATLAS, 2013); the larger the trial (sample size) the greater the statistical precision (Biau, Kernéis & Porcher, 2008). When comparing with trials in pharmacotherapy interventions for adult psychiatric disorder, trials in psychotherapy are less likely to have large sample sizes (and thus less statistical precision) and are also less likely to utilise control groups and masking (Huhn et al., 2014).

Depression has a relatively high prevalence in a population, suicide, however, is described as a "low base rate behavior", that is, there are relatively few events in a population (Brown & Jager-Hyman, 2014, p. 192). Trials of psychotherapy *specifically for* suicide prevention therefore require "very large" sample sizes to statistically reliably detect benefit (Brown & Jager-Hyman, 2014, p. 192).

Trial Quality and Bias

The measured beneficial effects of a range of psychotherapies in the treatment of adult depression were considered "probably overestimated because of publication bias and the relatively low quality of many studies in the field" in a comprehensive systematic review of studies, and thirty meta-analyses, from 1966-2010 (Cuijpers, Andersson, Donker & van Straten, 2011, p. 354). A review of the quality of psychotherapy trials for depression covering approximately the same period (1969-2011) did report a *trend* of improvement, though it could not be determined whether there were real improvements in quality, or improved reporting, or both (Chen et al., 2014).

The influence of biases in studies on the reported effects of psychotherapy for depression led Cuijpers, Karyotaki, Reijnders, & Ebert (2019, p. 21) to caution that the meta-analyses that do not adjust for bias "overestimate the effects considerably". For certain specific psychotherapies (including problem-solving therapy, interpersonal psychotherapy and BA) "insufficient evidence is available that they are effective" (Cuijpers et al., 2019). Specifically examining CBT for major depression disorder (MDD) and specified anxiety disorders, it was reported in the results of a comprehensive meta-analytic review that "the methodological quality in most of the studies was low or unknown", and there were some indications of publication bias (Cuijpers, Cristea, Karyotaki, Reijnders & Huibers, 2016, p. 253). Although it was concluded that CBT is *probably* effective for MDD (and the specified anxiety disorders), the authors noted that the effects are uncertain due to the "small number of high-quality trials", and that the results should therefore be "considered with caution" (Cuijpers et al., 2016, p. 254). Thus, when reviewing trial and meta-analysis results, as in other fields of medicine,

attention should be paid to the potential presence of bias, and the effect that this may have had on outcomes.

Use of Non-inferiority Test

A trial noted above (Richards et al, 2016) utilised a non-inferiority design, and concluded that BA, delivered by junior mental health workers, is no less effective for depression than CBT, for which highly-trained (and thus more costly) practitioners are required. In correspondence following the trial publication, concerns were raised concerning the use of the non-inferiority test. Vieta & Sanchez-Moreno (2017) noted that limitations, inherent in the use of a non-inferiority trial design, had not been mentioned; with no third (placebo) arm, against which both active arms could have demonstrated superiority, both therapies might have been equally *in*effective rather than effective. Vieta & Sanchez-Moreno considered that when assessing psychometric outcomes, only a trial utilising a superiority test can establish efficacy. Rief & Hofmann (2018) considered significant superiority results more reliable, and that aspects of the non-inferiority testing may tend towards an (incorrect) conclusion of non-inferiority.

The non-inferiority test (when used with no third placebo arm) is based on findings outside of the study - that the comparison therapy (the control) has the same treatment effect in the non-inferiority trial as it did in the forerunner (superiority) trials that established its effectiveness (FDA, 2016, cited in Brennan, 2016). Such study designs have similarities with (and therefore limitations of) historically-controlled trials (FDA, 2016); study participants in non-inferiority trial may not match the characteristics of those in the earlier superiority trials.

In contrast, a significant result in a trial testing *superiority* (treatment versus control), stands without reference to any earlier studies (FDA, 2016, cited in Brennan, 2016).

Exclusion Criteria

Trials necessarily have entry criteria, which can lead, for example, to the *exclusion* of patients with comorbidities and this may be a reason why outcomes in clinical practice may not be as good as trial results had indicated (Dodwell, Richmond, Hazell & Coleman, 2017).

In psychotherapy trials there appears to be a risk that if individuals at high risk of suicide are excluded from potentially practice-informing trials, the results may favour an intervention *less effective* for those most at risk. Included in the exclusion criteria for the non-inferiority study of BA versus CBT discussed above (Richards et al, 2016), were those at high risk of suicide. Baumeister (2017) in subsequent correspondence in the same journal, noted those eligible for enrolment may not be the patients that would be of greatest concern.

When considering the appropriateness of BA, against the more expensive CBT, it should be noted that there is evidence from randomised trials that CBT focusing on *suicidal cognitions and behaviours,* is effective (Mewton & Andrews,

2016). As noted by Baumeister (2017, p. 367) "the substitution of mental health care professionals by junior mental health workers does not consider the professional expertise needed in critical situations such as increased risk of suicide."

Concerns about exclusion criteria also apply to trials *directly* testing the effects of psychotherapy for suicide prevention; a lack of knowledge as to whether trial results indicating benefits of psychotherapies are valid in clinical practice has been reported (Brown & Jager-Hyman, 2014).

Use of Placebo Control

In medicine, the use of a placebo as a control in trials is "unlikely to affect underlying disease", however there is evidence placebos have "measurable effects" on, for example, depression (Blease, Bishop & Kaptchuk, 2017, p. 2). As a result of active participation in a study, those randomised to pill placebo have the frequent attention of professionals, and are provided with "positive expectations", "encouragement and support", "which may be sufficient to produce improvement" (Rief et al. 2009, cited in Cuijpers et al., 2014, p. 693).

In the pharmacotherapy field, the response to placebo has been found to be graduated according to the level of depression in the results of a meta-analysis of complete datasets of drug versus placebo trials submitted for the US licensing of a new generation antidepressants (Kirsch et al., 2008). A "substantial response to placebo" was reported for "moderately depressed" groups, however in those "with very severe levels of depression" the response diminished to an extent (Kirsch et al., 2008, p. 266), therefore it appears that the 'placebo effect' works less well with the most severely depressed.

In evaluating psychotherapy for depression, the use of a pill placebo as control can lead to a reduction in measured treatment effect, as the response in a pill placebo control group is likely to be greater than in a control group of individuals receiving no attention, i.e. 'no treatment'. The difference in treatment effect between psychotherapy and pill placebo was reported to be *substantially smaller* than that between psychotherapy and no treatment or waiting list control, in the results of a meta-analysis of trials comparing twelve psychotherapies with pill placebo for depression (Cuijpers et al., 2014). Trials with a pill placebo (or 'care-as-usual') control group yielded significantly lower measured treatment effect sizes than those with a waiting list control group in the results of the aforementioned meta-analytic review of CBT for MDD and specified anxiety conditions (Cuijpers et al, 2016).

Thus, in trials in assessing the effectiveness of treatments for depression, the use of pill placebo cannot be regarded as 'no treatment'.

Length of Follow-up

Depression can be a long-term condition with high relapse rates, thus "…whether treatments are equally efficacious for the longer term is at least as important as their efficacy in the short term" (Cuijpers, van Straten, Andersson &

van Oppen, 2008, p. 919). In evaluating the relative effectiveness of pharmacological versus psychotherapy interventions, it should be noted that there are indications that effects of the latter may "last longer" than those of the former (Cuijpers et al., 2014, p. 693). Any "strong effects" of antidepressants cease when they are no longer taken, whilst psychological therapies "may have sustained effects over the longer-term" (Cuijpers, et al., 2013, p. 146).

Published data meta-analyses are dependent on the length of participant follow-up as reported by the trialists. The use of individual participant data (IPD) meta-analyses provide more reliable results as the data can be checked at participant level, and the long-term follow-up of trial participants can be enabled, and thus any delayed effects of a treatment established (Riley, Lambert & Ghada Abo-Zaid, 2010; EBCTCG, 2005). In medicine, for example cancer treatment research, the results of IPD MAs have been practice changing since the mid-1980s, with participant follow-up that can extend to 15 years or more (EBCTCG, 2005). In the psychotherapy field, the IPD method of meta-analysis has been utilised less often, however, in the past decade some IPD MAs have been published, or are underway (Ebert et al., 2018). The increasing use of IPD MAs in this field can lead to more reliable and accurate meta-analyses and the potential for longer term follow-up.

Caution in the Interpretation of Subgroup Results

A meta-analysis of trial data provides greater statistical reliability than the individual trials as a result of pooling the data. In a protocol for an IPD MA in psychological interventions in pre-clinical depression it was noted that the use of individual participant data may enable the estimation of "specific effects" in subgroups (Ebert et al, 2018, p. 5-6). Although it is possible to reliably analyse subgroups, the selecting of subgroups is moving away from the larger dataset created by the meta-analysis, and the benefit of being able to *reliably* detect small to moderate treatment effects (Baigent, 1997) may be lost.

Whether in a large single trial or in a meta-analysis of trials, subgroups *inevitably* have less randomised data (than the complete dataset) and thus there is a greater potential for the effect of the 'play of chance' in subgroups than in the overall result (Cuzick, 2005). There are "statistical difficulties" in determining whether "…a treatment works better in some subgroups than others", which "has led to much misinterpretation" (Cuzick, 2005, p. 1308).

Multiple comparisons in subgroups can lead to false positives, and inadequate statistical power (insufficient data) can lead to false negatives, and it has been considered "uncertain when subgroup analyses should influence clinical practice", though statistical guidance is available to improve analyses (Burke, Sussman, Kent & Hayward, 2015). Thus, although the analysis of subgroups can enable the deduction of useful additional findings, care is required in the statistical interpretation.

Conclusions

A selection of aspects of trials that may influence the applicability of trial results to practice have been highlighted. There are inevitably limitations inherent in any given trial, and a report may, or may not, draw a reader's attention to these. Limitations of trial designs *in general* have been discussed, but also *specific* issues with the application of this method (of evaluating effectiveness) to psychotherapy for depression, together with the interrelated issue of risk of suicide.

The application of randomised controlled trials to psychotherapy, to provide reliable evidence for best practice, was later than in some other areas of medicine; the trials were often of small size and/or of poor quality, and with indications of publication bias.

Established general concerns with the use of the non-inferiority trial design apply also to the application of this methodology to psychotherapy. Trial results, when there is only a two-arm randomisation with no control arm to verify treatment efficacy with the specific trial participants selected, should be regarded with caution.

All trials necessarily have eligibility criteria, the specific exclusion of those at high risk of suicide from trial(s) evaluating psychological interventions for depression, appears to create a risk that the implementation of a treatment practice, informed by trial results, may envelop the category of individuals that were excluded from entry into the practice informing trial(s).

Meta-analyses in psychotherapy for depression have predominantly relied on published data. The more reliable, individual patient data method, which can facilitate long-term follow-up, only being utilised in relatively recent years. Large datasets, such as those created by an IPD meta-analysis, can enable subgroup analyses, however there are potential hazards with such analyses, as the statistical benefits (large numbers) of the *complete* meta-analysis dataset are diminished, and there is greater potential for the 'play of chance', and incorrect inferences drawn**.**

The establishment of well-evidenced and effective therapies for depression could lead to a marked reduction in the prevalence and societal burden of this condition.

Recommendations:

- When evaluating trial results, the interrelationship between depression, all-cause mortality and risk of suicide should be borne in mind.
- There is greater confidence in trial results when a robust trial methodology has been demonstrably utilised (Zalsman et al., 2016).
- There are potential benefits in the long-term follow-up of trial participants.

- Attention should be paid to ensuring treatments are not inappropriately applied to categories of individuals excluded from practice-informing trials.
- Care is needed in assessing the reliability of subgroup analyses.
- Where a non-inferiority design is used, a third placebo arm is advisable, against which both active arms can demonstrate superiority.
- In trials involving participants with depression, a pill placebo may not be considered 'no treatment' and thus the use thereof may reduce measured treatment effect.
- The use of largest practicable sample size maximises the statistical precision.

References

Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) Collaborative Group. (2013). Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *The Lancet, 381*, 805–816. doi: 10.1016/S0140-6736(12)61963-1

Baigent, C. (1997). The need for large-scale randomized evidence. *British Journal of Clinical Pharmacology, 43*, 349-353.

Barkham, M., Moller, N. P., & Pybis, J. (2017). How should we evaluate research on counselling and the treatment of depression? A case study on how the National Institute for Health and Care Excellence's draft 2018 guideline for depression considered what counts as best evidence. *Counselling and Psychotherapy Research, 17*(4), 253-268. doi: 10.1002/capr.12141

Baumeister, H. (2017). Behavioural activation training for depression. Letter. The *Lancet, 389*, 366-367.

Bernal, M., Haro, J. M., Bernert, S., Brugha, T., de Graaf, R., Bruffaerts, R., . . . Alonso, J. (2007). Risk factors for suicidality in Europe: results from the ESEMED study. *Journal of Affective Disorders, 101*, 27-34. doi: 10.1016/j.jad.2006.09.018

Biau, D. J., Kernéis, S., & Porcher, R. (2008). Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clinical orthopaedics and related research*, *466*(9), 2282–2288. doi:10.1007/s11999-008-0346-9

Blease, C. R., Bishop, F. L., & Kaptchuk, T. J. (2017). Informed consent and clinical trials: where is the placebo effect? *British Medical Journal, 356*. doi: 10.1136/bmj.j463

Brennan, Z. (2016). When can non-inferiority trials establish efficacy? FDA explains with guidance. *Regulatory Focus*. Posted 07 November 2016. Regulatory Affairs Professionals Society (RAPS). Maryland: RAPS. Retrieved from https://www.raps.org/ regulatory-focus™/news-articles/2016/11/when-can-non-inferiority-trials-establish-efficacy-fda-explains-with-guidance

Brown, G. K., & Jager-Hyman, S. (2014). Evidence-Based Psychotherapies for Suicide Prevention: Future Directions. *American Journal of Preventive Medicine, 47*(3), Suppl. 2, 186-194. doi: 10.1016/j.amepre.2014.06.008.

Burke, J. F., Sussman, J. B., Kent, D. M., & Hayward, R. A. (2015). Three simple rules to ensure reasonably credible subgroup analyses *British Medical Journal, 351*:h5651. doi.org/10.1136/bmj.h5651

Button, K. S., Ioannidis, J. P. A., Mokrysz C., Nosek B. A., Flint J., Robinson E. S. J., & Munafò M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews: Neuroscience, 14*, 365-376. doi:10.1038/nrn3475

Chen, P., Furukawa, T. A., Shinohara, K., Honyashiki, M., Imai, H., Ichikawa, K., . . . Churchill, R. (2014). Quantity and quality of psychotherapy trials for depression in the past five decades. *Journal of Affective Disorders, 165*, 190-195. doi: 10.1016/ j.jad.2014.04.071

Chesney, E., Goodwin, G. M. & Fazel, S. (2014), Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry, 13*: 153-160. doi:10.1002/wps.20128

Cooper, M. (2011). *Briefing paper: development of a randomised controlled trial of counselling for depression*. British Association for Counselling & Psychotherapy (BACP). Final draft 23 March 2011. Lutterworth, UK: BACP.

Crofton, J. (2006). The MRC randomized trial of streptomycin and its legacy: a view from the clinical front line. *Journal of the Royal Society of Medicine, 99*(10), 531-534. doi: 10.1258/jrsm.99.10.531

Cuijpers P., van Straten A., Andersson G., & van Oppen P. (2008). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *J Consult Clin Psychol. 76*(6), 909-922. doi: 10.1037/a0013075.

Cuijpers, P., Andersson, G., Donker, T., & van Straten, A. (2011). Psychological treatment of depression: results of a series of meta-analyses. *Nordic Journal of Psychiatry, 65*(6), 354-364. doi: 10.3109/08039488.2011.596570

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. *World Psychiatry, 12*: 137-148. doi:10.1002/wps.20038

Cuijpers, P., Turner, E. H., Mohr, D. C., Hofmann, S. G., Andersson, G., Berking, M., & Coyne, J. (2014). Comparison of psychotherapies for adult depression to pill placebo groups: a meta-analysis. *Psychological Medicine, 44*, 685-695. doi: 10.1017/S003329 1713000457

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry, 15*, 245-258. doi: 10.1002/wps.20346

Cuijpers, P., Karyotaki, E., Reijnders, M., & Ebert, D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences, 28*(1), 21-30. doi:10.1017/S2045796018000057

Cuzick, J., (2005). Forest plots and the interpretation of subgroups. *The Lancet, 365*, 1308.

Dhossche, D. M., Meloukheia, A. M. & Chakravorty, S. (2000). The association of suicide attempts and comorbid depression and substance abuse in psychiatric consultation patients. *General Hospital Psychiatry, 22*, 281-288.

Diagnostic and Statistical Manual of Mental Disorders, DSM-5. (2018). American Psychiatric Association (ACA). Washington DC: ACA. https://dsm.psychiatryonline.org/pb-assets/dsm/update/DSM5Update_October2018.pdf

Dodwell, D., Richmond, A., Hazell, S., & Coleman, R. Doing more with less; are there some 'easy wins' in the management of early breast cancer? (2017). *Clinical Oncology, 29*(9), 550-552. doi.org/10.1016/j.clon.2017.04.033

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2005). Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. *The Lancet, 366*(9503), 2087-2106. doi.org/10.1016/S0140-6736(05)67887-7

Ebert, D. D., Buntrock, C., Reins, J. A., Zimmermann, J. & Cuijpers, P. (2018). Efficacy and moderators of psychological interventions in treating subclinical symptoms of depression and preventing major depressive disorder onsets: protocol for an individual patient data meta-analysis of randomised controlled trials. *BMJ Open, 8:* e018582. doi:10.1136/bmjopen-2017-018582

Gilman, S. E., Sucha, E., Kingsbury, M., Horton, N. J., Murphy, J. M., & Colman, I. (2017). Depression and mortality in a longitudinal study: 1952–2011. *Canadian Medical Association Journal, 189*(42), E1304-E1310. doi: 10.1503/cmaj.170125

Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Förstl, H., Pitschel-Walz, G., . . . Leucht, S. (2014). *JAMA Psychiatry, 71*(6), 706-715. doi: 10.1001/jamapsychiatry.2014.112

Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J. & Johnson, B. T. (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine, 5,* 250-268. doi: 10.1371/journal. pmed.0050045

McDonald, S., Westby, M., Clarke, M. & Lefebvre, C. (2002). Number and size of randomized trials reported in general health care journals from 1948 to 1997. *International Journal of Epidemiology, 31*(1), 125-127.

Medical Research Council (MRC). (1948). Medical Research Council Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment for pulmonary tuberculosis. *British Medical Journal, ii*, 769-782.

Méndez-Bustos, P., Calati, R., Rubio-Ramírez, F., Olié, E., Courtet, P., & Lopez-Castroman, J. (2019). *Frontiers in Psychology, 10*:277. doi:10.3389/fpsyg.2019.00277

Mewton, L., & Andrews, G. (2016). Cognitive behavioural therapy for suicidal behaviors: improving patient outcomes. *Psychology Research and Behavior Management, 9,* 21-29. doi: 10.2147/PRBM.S84589

Miret, M., Ayuso-Mateos, J. L., Sanchez-Moreno, J., & Vieta, E. (2013). Depressive disorders and suicide: epidemiology, risk factors, and burden. *Neuroscience & Biobehavioral Reviews, 37*(10), 2372-2374. doi: 10.1016/j.neubiorev.2013.01.008.

National Institute for Health Care Excellence (NICE) (2018). Retrieved from www.nice.org.uk

National Institute for Health Care Excellence (NICE) 'Depression in adults: recognition and management'. Clinical guideline [CG90]. 2009, last updated April 2018. Retrieved 3 March 2019 from https://www.nice.org.uk/guidance/cg90

Office for National Statistics (ONS). (2017). Suicides in the UK. (London: ONS). Retrieved from www.ons.gov.uk

Pearson, A., Saini, P., Da Cruz, D., Miles, C., While, D., Swinson, N., . . . Kapur, N. (2009). Primary care contact prior to suicide in individuals with mental illness. *British Journal of General Practice, 59*, 825-832. doi: 10.3399/bjgp09X472881

Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., . . . Finning, K. (2016). Cost and outcome of Behavioural Activation versus Cognitive Behavioural Therapy for depression (COBRA): a randomised, controlled, non-inferiority trial. *The Lancet, 388*, 871-880. doi: 10.1016/S0140-6736(16)31140-0

Richards D. A., Hill J. J., Gask L., Lovell K., Chew-Graham C., Bower P., . . . Barkham M. (2013). Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial *British Medical Journal, 347*:f4913. doi: 10.1136/bmj.f4913

Rief, W., & Hofmann, S. G. (2018). Some problems with non-inferiority tests in psycho-therapy research: psychodynamic therapies as an example. *Psychological Medicine, 48*(8), 1392-1394. doi: 10.1017/S0033291718000247

Riley, R. D., Lambert, P. C. & Ghada Abo-Zaid (2010). Meta-analysis of individual participant data: rationale, conduct, and reporting. *British Medical Journal, 340*, c221. doi: 10.1136/bmj.c221.

Saxon D., Ashley K., Bishop-Edwards L., Connell J., Harrison P., Ohlsen S., … Barkham M. (2017). A pragmatic randomised controlled trial assessing the non-inferiority of counselling for depression versus cognitive-behaviour therapy for patients in primary care meeting a diagnosis of moderate or severe depression (PRaCTICED): Study protocol for a randomised controlled trial. *Trials. 18*:93. doi: 10.1186/s13063-017-1834-6

Schneider, B. (2012). Behavioural therapy of suicidality. *European Archives of Psychiatry and Clinical Neuroscience, 262*(Suppl. 2), 123. doi:10.1007/s00406-012-0351-6

Second International Study of Infarct Survival Collaborative Group (ISIS-2). (1988). Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction. *The Lancet, 332*(8607), 349-360.

U.S. Food & Drug Administration (FDA). (2016). *Non-Inferiority Clinical Trials to Establish Effectiveness. Guidance for Industry*. November 2016. Maryland: FDA.

Vieta, E., & Sanchez-Moreno, J. (2017). Behavioural activation training for depression. Letter. *The Lancet, 389*:367.

World Health Organization (WHO). (2014). Preventing suicide, a global imperative. Executive summary. WHO/MSD/MER/14.2. (Zurich: WHO).

Zalsman, G., Hawton, K., Wasserman, D., van Heeringen, K., Arensman, E., Sarchiapone, M., . . . Zohar, J. (2016). Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry, 3*(7), 646-659. doi: 10.1016/S2215-0366(16)30030-X